

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
25 March 2004 (25.03.2004)

PCT

(10) International Publication Number  
**WO 2004/024886 A2**

- (51) International Patent Classification<sup>7</sup>: C12N (72) Inventor: OLEINIKOV, Andrew, V.; 16208 32nd Ave. SE, Mill Creek, WA 98012 (US).
- (21) International Application Number: PCT/US2003/028946 (74) Agent: OSTER, Jeffrey, B.; Combimatrix Corporation, 6500 Harbour Heights Parkway, Mukilteo, WA 98275 (US).
- (22) International Filing Date: 12 September 2003 (12.09.2003) (81) Designated States (*national*): AU, CA, IL, JP, NO, NZ, RU.
- (25) Filing Language: English (84) Designated States (*regional*): European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).
- (26) Publication Language: English
- (30) Priority Data: 10/243,367 12 September 2002 (12.09.2002) US Published:  
— without international search report and to be republished upon receipt of that report
- (71) Applicant: COMBIMATRIX CORPORATION [US/US]; 6500 Harbour Heights Parkway, Mukilteo, WA 98275 (US). For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: MICROARRAY SYNTHESIS AND ASSEMBLY OF GENE-LENGTH POLYNUCLEOTIDES

(57) Abstract: There is disclosed a process for *in vitro* synthesis and assembly of long, gene-length polynucleotides based upon assembly of multiple shorter oligonucleotides synthesized *in situ* on a microarray platform. Specifically, there is disclosed a process for *in situ* synthesis of oligonucleotide fragments on a solid phase microarray platform and subsequent, "on device" assembly of larger polynucleotides composed of a plurality of shorter oligonucleotide fragments.

WO 2004/024886 A2

## MICROARRAY SYNTHESIS AND ASSEMBLY OF GENE-LENGTH POLYNUCLEOTIDES

### Technical Field of the Invention

5       The present invention provides a process for *in vitro* synthesis and assembly of long, gene-length polynucleotides based upon assembly of multiple shorter oligonucleotides synthesized *in situ* on a microarray platform. Specifically, the present invention provides a process for *in situ* synthesis of oligonucleotide sequence fragments on a solid phase microarray platform and subsequent, "on chip" assembly of larger polynucleotides composed of a plurality  
10 of smaller oligonucleotide sequence fragments.

### Background of the Invention

In the world of microarrays, biological molecules (*e.g.*, oligonucleotides, polypeptides and the like) are placed onto surfaces at defined locations for potential binding with target samples of nucleotides or receptors. Microarrays are miniaturized arrays of biomolecules  
15 available or being developed on a variety of platforms. Much of the initial focus for these microarrays have been in genomics with an emphasis of single nucleotide polymorphisms (SNPs) and genomic DNA detection/validation, functional genomics and proteomics (Wilgenbus and Lichter, *J. Mol. Med.* 77:761, 1999; Ashfari et al., *Cancer Res.* 59:4759, 1999; Kurian et al., *J. Pathol.* 187:267, 1999; Hacia, *Nature Genetics* 21 suppl.:42, 1999; Hacia et  
20 al., *Mol. Psychiatry* 3:483, 1998; and Johnson, *Curr. Biol.* 26:R171, 1998).

There are, in general, three categories of microarrays (also called "biochips" and "DNA Arrays" and "Gene Chips" but this descriptive name has been attempted to be a trademark) having oligonucleotide content. Most often, the oligonucleotide microarrays have a solid surface, usually silicon-based and most often a glass microscopic slide. Oligonucleotide  
25 microarrays are often made by different techniques, including (1) "spotting" by depositing single nucleotides for *in situ* synthesis or completed oligonucleotides by physical means (ink jet printing and the like), (2) photolithographic techniques for *in situ* oligonucleotide synthesis (see, for example, Fodor U.S. Patent '934 and the additional patents that claim priority from this priority document, (3) electrochemical *in situ* synthesis based upon pH based removal of  
30 blocking chemical functional groups (see, for example, Montgomery U.S. Patent 6,092,302 the disclosure of which is incorporated by reference herein and Southern U.S. Patent 5,667,667), and (4) electric field attraction/repulsion of fully-formed oligonucleotides (see, for example, Hollis et al., U.S. Patent 5,653,939 and its duplicate Heller U.S. Patent 5,929,208). Only the first three basic techniques can form oligonucleotides *in situ* that are, building each  
35 oligonucleotide, nucleotide-by-nucleotide, on the microarray surface without placing or attracting fully formed oligonucleotides.

With regard to placing fully formed oligonucleotides at specific locations, various micro-spotting techniques using computer-controlled plotters or even ink-jet printers have been developed to spot oligonucleotides at defined locations. One techniques loads glass fibers

having multiple capillaries drilled through then with different oligonucleotides loaded into each capillary tube. Microarray chips, often simply glass microscope slides, are then stamped out much like a rubber stamp on each sheet of paper or glass slide. It is also possible to use "spotting" techniques to build oligonucleotides *in situ*. Essentially, this involves "spotting" relevant single nucleotides at the exact location or region on a slide (preferably a glass slide) where a particular sequence of oligonucleotide is to be built. Therefore, irrespective of whether or not fully formed oligonucleotides or single nucleotides are added for *in situ* synthesis, spotting techniques involve the precise placement of materials at specific sites or regions using automated techniques.

Another technique involves a photolithography process involving photomasks to build oligonucleotides *in situ*, base-by-base, by providing a series of precise photomasks coordinated with single nucleotide bases having light-cleavable blocking groups. This technique is described in Fodor et al., U.S. Patent 5,445,934 and its various progeny patents. Essentially, this technique provides for "solid-phase chemistry, photolabile protecting groups, and photolithography . . . to achieve light-directed spatially-addressable parallel chemical synthesis."

The electrochemistry platform (Montgomery U.S. Patent 6,092,302, the disclosure of which is incorporated by reference herein) provides a microarray based upon a semiconductor chip platform having a plurality of microelectrodes. This chip design uses Complimentary Metal Oxide Semiconductor (CMOS) technology to create high-density arrays of microelectrodes with parallel addressing for selecting and controlling individual microelectrodes within the array. The electrodes turned on with current flow generate electrochemical reagents (particularly acidic protons) to alter the pH in a small "virtual flask" region or volume adjacent to the electrode. The microarray is coated with a porous matrix for a reaction layer material. Thickness and porosity of the material is carefully controlled and biomolecules are synthesized within volumes of the porous matrix whose pH has been altered through controlled diffusion of protons generated electrochemically and whose diffusion is limited by diffusion coefficients and the buffering capacities of solutions. However, in order to function properly, the microarray biochips using electrochemistry means for *in situ* synthesis has to alternate anodes and cathodes in the array in order to generate needed protons (acids) at the anodes so that the protons and other acidic electrochemically generated acidic reagents will cause an acid pH shift and remove a blocking group from a growing oligomer.

#### Gene Assembly

The preparation of arbitrary polynucleotide sequences is useful in a "post-genomic" era because it provides any desirable gene oligonucleotide or its fragment, or even whole genome material of plasmids, phages and viruses. Such polynucleotides are long, such as excess of 1000 bases in length. *In vitro* synthesis of oligonucleotides (given even the best yield conditions of phosphoramidite chemistry) would not be feasible because each base addition reaction is less than 100% yield. Therefore, researchers desiring to obtain long

polynucleotides of gene length or longer had to turn to nature or gene isolation techniques to obtain polynucleotides of such length. For the purposes of this patent application, the term "polynucleotide" shall be used to refer to nucleic acids (either single stranded or double stranded) that are sufficiently long so as to be practically not feasible to make *in vitro* through single base addition. In view of the exponential drop-off in yields from nucleic acid synthesis chemistries, such as phosphoramidite chemistry, such polynucleotides generally have greater than 100 bases and often greater than 200 bases in length. It should be noted that many commercially useful gene cDNA's often have lengths in excess of 1000 bases.

Moreover, the term "oligonucleotides" or shorter term "oligos" shall be used to refer to shorter length single stranded or double stranded nucleic acids capable of *in vitro* synthesis and generally shorter than 150 bases in length. While it is theoretically possible to synthesize polynucleotides through single base addition, the yield losses made it a practical impossibility beyond 150 bases and certainly longer than 250 bases.

However, knowledge of the precise structure of the genetic material is often not sufficient to obtain this material from natural sources. Mature cDNA, which is a copy of an mRNA molecule, can be obtained if starting material contains this desired mRNA. However, it is not always known if this particular mRNA is present in a sample or the amount of this mRNA might be too low to obtain the corresponding cDNA without significant difficulties. Also, different levels of homology or splice variants may interfere with obtaining one particular species of mRNA. On the other hand many genomic materials might be not appropriate to prepare mature gene (cDNA) due to exon-intron structure of genes in many different genomes.

In addition, there is a need in the art for polynucleotides not existing in nature to improve genomic research performance. In general, the ability to obtain polynucleotide of any desired sequence just knowing the primary structure for the reasonable price in short period of time will significantly move forward several fields of biomedical research and clinical practice.

Assembly of long arbitrary polynucleotides from oligonucleotides synthesized by organic synthesis and individually purified has other problems. The assembly can be performed using PCR or ligation methods. The synthesis and purification of many different oligonucleotides by conventional methods (even using multi-channel synthesizers) are laborious and expensive procedures. The current price of assembled polynucleotide on the market is about \$12-25 per base pair, which can be considerable for assembling larger polynucleotides. Very often the amount of conventionally synthesized oligonucleotides would be excessive. This also contributes to the cost of the final product.

Therefore, there is a need in the art to provide cost-effective polynucleotides by procedures that are not as cumbersome and labor-intensive as present methods to be able to provide polynucleotides at costs below \$1 per base or 1-20 times less than current methods. The present invention was made to address this need.

### Summary of the Invention

The present invention provides a process for the assembly of oligonucleotides synthesized on microarrays into a polynucleotide sequence. The desired target polynucleotide sequence is dissected into pieces of overlapping oligonucleotides. In the first embodiment these oligonucleotides are synthesized *in situ*, in parallel on a microarray chip in a non-cleavable form. A primer extension process assembles the target polynucleotides. The primer extension process uses starting primers that are specific for the appropriate sequences. The last step is PCR amplification of the final polynucleotide product. Preferably, the polynucleotide product is a cDNA suitable for transcription purposes and further comprising a promoter sequence for transcription.

The present invention provides a process for assembling a polynucleotide from a plurality of oligonucleotides comprising:

- (a) synthesizing or spotting a plurality of oligonucleotide sequences on a microarray device or bead device having a solid or porous surface, wherein a first oligonucleotide is oligo 1 and a second oligonucleotide is oligo 2 and so on, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein the first oligonucleotide sequence has an overlapping sequence region of from about 10 to about 50 bases that is the same or substantially the same as a region of a second oligonucleotide sequence, and wherein the second oligonucleotide sequence has an overlapping region with a third oligonucleotide sequence and so on;
- (b) forming complimentary oligo 1 by extending primer 1, wherein primer 1 is complimentary to oligo 1;
- (c) disassociating complimentary oligo 1 from oligo 1 and annealing complimentary oligo 1 to both oligo 1 and to the overlapping region of oligo 2, wherein the annealing of complimentary oligo 1 to oligo 2 serves as a primer for extension for forming complimentary oligo 1+2;
- (d) repeating the primer extension cycles of step (c) until a full-length polynucleotide is produced; and
- (e) amplifying the assembled complementary full length polynucleotide to produce a full length polynucleotide in desired quantities.

Preferably, the solid or porous surface is in the form of a microarray device. Most preferably, the microarray device is a semiconductor device having a plurality of electrodes for synthesizing oligonucleotides *in situ* using electrochemical means to couple and decouple nucleotide bases. Preferably, the primer extension reaction is conducted through a sequential process of melting, annealing and then extension. Most preferably, the primer extension reaction is conducted in a PCR amplification device using the microarray having the plurality of oligonucleotides bound thereto.

The present invention further provides a process for assembling a polynucleotide from a plurality of oligonucleotides comprising:

- (a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence has an overlapping region corresponding to a next oligonucleotide sequence within the sequence and further comprises two flanking sequences, one at the 3' end and the other at the 5' end of each oligonucleotide, wherein each flanking sequence is from about 7 to about 50 bases and comprising a primer region and a sequence segment having a restriction enzyme cleavable site;
- (b) amplifying each oligonucleotide using the primer regions of the flanking sequence to form double stranded (ds) oligonucleotides;
- (c) cleaving the oligonucleotide sequences at the restriction enzyme cleavable site; and
- (d) assembling the cleaved oligonucleotide sequences through the overlapping regions to form a full length polynucleotide.
- 15 Preferably, the flanking sequence is from about 10 to about 20 bases in length. Preferably, the restriction enzyme cleavable site is a class II endonuclease restriction site sequence capable of being cleaved by its corresponding class II restriction endonuclease enzyme. Most preferably, the restriction endonuclease class II site corresponds to restriction sites for a restriction endonuclease class II enzyme selected from the group consisting of Mly I, BspM I, Bae I, BsaX I, Bsr I, Bmr I, Btr I, Bts I, Fok I, and combinations thereof. Preferably, the flanking sequence further comprises a binding moiety used to purify cleaved oligonucleotides from flanking sequences. Preferably, the process further comprises the step of labeling the flanking sequence during the amplification step (b) using primer sequences labeled with binding moieties. Most preferably, a binding moiety is a small molecule able to be captured, such as biotin captured by avidin or streptavidin, or fluorescein able to be captured by an anti-fluorescein antibody.

The present invention further provides a process for assembling a polynucleotide from a plurality of oligonucleotides comprising:

- (a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence has an overlapping region corresponding to a next oligonucleotide sequence within the sequence, and further comprises a sequence segment having a cleavable linker moiety;
- (b) cleaving the oligonucleotide sequences at the cleavable linker site to cleave each oligonucleotide complex from the microarray or bead solid surface to form a soluble mixture of oligonucleotides, each having an overlapping sequence; and
- (c) assembling the oligonucleotide sequences through the overlapping regions to form a full length polynucleotide.

Preferably, the cleavable linker is a chemical composition having a succinate moiety bound to a nucleotide moiety such that cleavage produces a 3'hydroxy nucleotide. Most preferably, the cleavable linker is selected from the group consisting of 5'-dimethoxytrityl-thymidine-3'-succinate, 4-N-benzoyl-5'-dimethoxytrityl-deoxycytidine-3'-succinate, 1-N-benzoyl-5'-dimethoxytrityl-deoxyadenosine-3'-succinate, 2-N-isobutyryl-5'-dimethoxytrityl-deoxyguanosone-3'-succinate, and combinations thereof.

The present invention further provides a process for assembling a polynucleotide from a plurality of oligonucleotides comprising:

- (a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence has a flanking region at an end attached to the solid or porous surface, and a specific region designed by dissecting the polynucleotide sequence into a plurality of overlapping oligonucleotides, wherein a first overlapping sequence on a first oligonucleotide corresponds to a second overlapping sequence of a second oligonucleotide, and wherein the flanking sequence comprises a sequence segment having a restriction endonuclease (RE) recognition sequence capable of being cleaved by a corresponding RE enzyme;
- (b) hybridizing an oligonucleotide sequence complementary to the flanking region to form a double stranded sequence capable of interacting with the corresponding RE enzyme;
- (c) digesting the plurality of oligonucleotides to cleave them from the microarray device or beads into a solution; and
- (d) assembling the oligonucleotide mixture through the overlapping regions to form a full length polynucleotide.

Preferably, the flanking sequence is from about 10 to about 20 bases in length. Preferably, the restriction enzyme cleavable site is a class II endonuclease restriction site sequence capable of being cleaved by its corresponding class II restriction endonuclease enzyme. Most preferably, the restriction endonuclease class II site corresponds to restriction sites for a restriction endonuclease class II enzyme selected from the group consisting of Mly I, BspM I, Bae I, BsaX I, Bsr I, Bmr I, Btr I, Bts I, Fok I, and combinations thereof. Preferably, the process further comprises a final step of amplifying the polynucleotide sequence using primers located at both ends of the polynucleotide.

The present invention further provides a process for creating a mixture of oligonucleotide sequences in solution comprising:

- (a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence further comprises two flanking sequences, one at the 3' end and the other at the 5' end of each oligonucleotide, wherein each flanking sequence is from about 7 to

about 50 bases and comprising a primer region and a sequence segment having a restriction enzyme cleavable site;

(b) amplifying each oligonucleotide using the primer regions of the flanking sequence to form a double stranded (ds) oligonucleotides; and

5 (c) cleaving the double stranded oligonucleotide sequences at the restriction enzyme cleavable site.

Preferably, the flanking sequence is from about 10 to about 20 bases in length.

Preferably, the restriction enzyme cleavable site is a class II endonuclease restriction site sequence capable of being cleaved by its corresponding class II restriction endonuclease enzyme. Most preferably, the restriction endonuclease class II site corresponds to restriction 10 sites for a restriction endonuclease class II enzyme selected from the group consisting of Mly I, BspM I, Bae I, BsaX I, Bsr I, Bmr I, Btr I, Bts I, Fok I, and combinations thereof. Preferably, the flanking sequence further comprises a binding moiety used to purify cleaved oligonucleotides from flanking sequences. Preferably, the process further comprises the step of labeling the flanking sequence during the amplification step (b) using primer sequences 15 labeled with binding moieties. Most preferably, a binding moiety is a small molecule able to be captured, such as biotin captured by avidin or streptavidin, or fluorescein able to be captured by an anti-fluorescein antibody.

The present invention further provides a process for creating a mixture of 20 oligonucleotide sequences in solution comprising:

(a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence has a sequence segment having a cleavable linker moiety;

25 (b) cleaving the oligonucleotide sequences at the cleavable linker site to cleave each oligonucleotide sequence from the microarray or bead solid surface to form a soluble mixture of oligonucleotides.

Preferably, the cleavable linker is a chemical composition having a succinate moiety bound to a nucleotide moiety such that cleavage produces a 3'-hydroxy nucleotide. Most 30 preferably, the cleavable linker is selected from the group consisting of 5'-dimethoxytrityl-thymidine-3'-succinate, 4-N-benzoyl-5'-dimethoxytrityl-deoxycytidine-3'-succinate, 1-N-benzoyl-5'-dimethoxytrityl-deoxyadenosine-3'-succinate, 2-N-isobutyryl-5'-dimethoxytrityl-deoxyguanosone-3'-succinate, and combinations thereof.

The present invention further provides a process for creating a mixture of 35 oligonucleotide sequences in solution comprising:

(a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence has a flanking region at an end attached to the solid or porous



surface, and a specific region, wherein the flanking sequence comprises a sequence segment having a restriction endonuclease (RE) recognition sequence capable of being cleaved by a corresponding RE enzyme;

- (b) hybridizing an oligonucleotide sequence complementary to the flanking region
- 5 to form a double stranded sequence capable of interacting with the corresponding RE enzyme;
- (c) digesting the plurality of oligonucleotides to cleave them from the microarray device or beads into a solution.

Preferably, the flanking sequence is from about 10 to about 20 bases in length.

- 10 Preferably, the restriction enzyme cleavable site is a class II endonuclease restriction site sequence capable of being cleaved by its corresponding class II restriction endonuclease enzyme. Most preferably, the restriction endonuclease class II site corresponds to restriction sites for a restriction endonuclease class II enzyme selected from the group consisting of Mly I, BspM I, Bae I, BsaX I, Bar I, Bmr I, Btr I, Bts I, Fok I, and combinations thereof.

#### Brief Description of the Drawings

- 15 Figure 1 shows a schematic of gene assembly on a microarray device surface or porous matrix. In Figure 1A, the target gene sequence is dissected into number of overlapping oligonucleotides. The 3' and 5' are the ends of the shown strand. Figure 1A also shows, relative to the target sequence, primer Pr1; extension product of primer Pr1, which is complementary oligonucleotide 1; and extension product of complimentary oligonucleotide 1,
- 20 which is complementary oligonucleotides 1+2. Figure 1B illustrates one embodiment of the initial steps of an assembly process. In step 1 of assembly, Primer Pr1 is annealed to oligonucleotide 1 and extended by appropriate polymerase enzyme into product complementary to oligonucleotide 1. The second step is melting, re-annealing and extension (*i.e.*, amplification) to lead to production of larger amount of Pr1 extension product
- 25 (complementary oligonucleotide 1), re-association of the complementary oligonucleotide 1 with oligonucleotide 1, and to annealing of the complementary oligonucleotide 1 with oligonucleotide 2 followed by its extension into product complementary to oligonucleotides 1+2. Figure 1C shows a continuation of the assembly process from Figure 1B. Specifically, step 3 of the process (*i.e.*, melting, re-annealing and extension) leads to the same products as
- 30 step 2 plus a product complementary to oligonucleotides 1+2+3. Cycles (steps) are repeated until a full-length complementary polynucleotide is formed. The final step is preparation of the final target polynucleotide molecule in desirable amounts by amplification (*i.e.*, PCR) using two primers complementary to the ends of this molecule (PrX and PrY).

- 35 Figure 2 shows a second embodiment of the inventive gene assembly process using oligonucleotides synthesized *in situ* onto a microarray device, each having a flanking sequence region containing a restriction enzyme cleavage site, followed by a PCR amplification step and followed by a REII restriction enzyme cleavage step.

Figure 3 shows a schematic for gene assembly using oligos synthesized and then cleaved from a microarray device. Specifically, in the upper panel marked "A",

oligonucleotide sequences are connected to the microarray device through a cleavable linker (CL) moiety. An example of a cleavable linker moiety is provided in Figure 3A. The cleavable linkers are molecules that can withstand the oligonucleotide synthesis process (*i.e.*, phosphoramidite chemistry) and then can be cleaved to release oligonucleotide fragments.

- 5 Chemical cleavage at cleavable linker CL recreates usual 3' end of specific oligos 1 through N. These oligonucleotides are released into a mixture. The mixture of oligonucleotides is subsequently assembled into full-length polynucleotide molecules. In the lower panel marked "B" of Figure 3, oligonucleotide sequences are connected to the microarray device through additional flanking sequence containing a restriction enzyme (RE) sequence site. Another  
10 oligonucleotide sequence, complementary to the flanking sequence region, is hybridized to the oligonucleotides on the microarray device. This recreates a "ds" or double-stranded oligonucleotide structure, each having a RE sequence recognition region in the flanking sequence region. Digestion of this ds oligonucleotides with the corresponding RE enzymes at the RE recognition sites in the flanking sequence regions releases the specific oligonucleotides  
15 1 through N. When assembled, oligonucleotide sequences 1 through N form a full-length polynucleotide molecule.

- Figure 4 shows the assembly of a polynucleotide from three oligonucleotide fragments wherein each oligonucleotide fragment was synthesized *in situ* on a microarray device. The fully assembled polynucleotide was 172 mers in length, a length not practically achievable by  
20 *in situ* synthesis. The first embodiment inventive process was used in this example.

Figure 5 shows the oligonucleotide sequences used to assemble the 172-mer polynucleotide of Figure 4. The sequences of primers X and Z are underlined. The *Hpa* II restriction site is indicated by italic underlined letters.

- Figure 6 shows a scheme for preparing the sequences of flanking regions and primers  
25 used for preparation of specific oligonucleotide for assembly using the REII enzyme *Mly*I. Primer 1 is complementary to the oligonucleotide strand on a microarray device and contains a Biotin-TEG (triethylene glycol) moiety. Primer 2 is the same strand as the oligonucleotide strand on microarray device and contains Biotin-TEG moiety. Any sequence between the primers can be used and is just designated by a string of N's.

- 30 Figure 7 shows the results of PCR and *Mly*I digestion of an oligonucleotide sequence as described in Figure 6. The clean bands show the ability to obtain pure oligonucleotides using the second embodiment of the inventive process to cleave off oligonucleotide sequences using appropriate restriction enzymes.

- Figure 8 shows the sequences from nine oligonucleotides fragments (consecutively  
35 numbered 1-9) used to assemble a 290 bp polynucleotide. The flanking regions are shown in bold and underlined. The process used for polynucleotide assembly was the second embodiment. The overlapping regions further contained a cleavable site as the *Mly*I recognition site for the *Mly*I class II restriction endonuclease.

Figure 9 shows a schematic in the top panel for assembling a polynucleotide from nine oligonucleotides. Nine oligonucleotide sequences, shown in Figure 8, were amplified by PCR using primers 1 and 2 (as described in Figure 6) into ds DNA fragments containing the same flanking regions and specific overlapping sequences, digested with *Mly*I enzyme to remove flanking sequences, and used for assembly of 290 bp DNA fragment. The columns in the gel shown are M – markers, 1 – negative control, assembly without primers FP1 and FP2, 2 – negative control, assembly without specific oligos, 3 – assembly of 290 bp fragment from specific oligos plus amplification with FP1 and FP2 primers. The band in column 3 shows a high efficiency of the inventive polynucleotide assembly process.

Figure 10 shows a sequence of an assembled polynucleotide in Example 4, broken down into its component oligonucleotides.

#### Detailed Description of the Invention

The present invention describes the preparation of a polynucleotide sequence (also called “gene”) using assembly of overlapping shorter oligonucleotides synthesized or spotted on microarray devices or on solid surface bead devices. The shorter oligonucleotides include sequence regions having overlapping regions to assist in assembly into the sequence of the desired polynucleotide. By overlapping regions, it is referred to sequence regions at either a 3’ end or a 5’ end of a first oligonucleotide sequence that is the same as part of the second oligonucleotide and has the same direction (relative to 3’ to 5’ or 5’ to 3’ direction), and will hybridize to the 5’ end or 3’ end of a second oligonucleotide sequence or its complimentary sequence (second embodiment), and a second oligonucleotide sequence to third oligonucleotide sequence, and so on. In order to design or develop a microarray device or bead device to be used for polynucleotide assembly, the polynucleotide sequence is divided (or dissected) into a number of overlapping oligonucleotides segments, each with lengths preferably from 20 to 1000 bases, and most preferably from 20 to 200 bases (Figure 1A). The overlap between oligonucleotide segments is 5 or more bases, preferably 15 – 25 bases to that proper hybridization of first to second, second to third, third to fourth and so on occurs. These oligonucleotides (or oligos) are preferably synthesized on a microarray device using any available method (*i.e.*, electrochemical *in situ* synthesis, photolithography *in situ* synthesis, ink-jet printing, spotting, etc.). The direction of synthesis relative to the microarray device surface or porous matrix covering a microarray device can be from 3’ to 5’ or from 5’ to 3’. Preferably, *in situ* synthesis is done in the 3’ to 5’ direction.

In the first embodiment the inventive gene/polynucleotide assembly process uses oligonucleotides immobilized on a microarray device. The microarray device itself or a porous reaction layer with immobilized oligonucleotides can be used for the inventive gene/polynucleotide assembly process.

With regard to Figure 1B, the process comprises several repeated steps of melting, annealing and extension (Figure 1B), which can be performed in any thermal cycler instrument. The cycling program is similar to the programs used for PCR. At the first step of

gene/polynucleotide assembly, primer Pr1 is added and anneals to oligonucleotide 1 on the microarray device and then extends by appropriate polymerase enzyme into product complementary to oligonucleotide 1 (called complimentary oligonucleotide 1). At the second step of the process the product complementary to oligonucleotide 1 is melted from  
5 oligonucleotide 1, primer Pr1 is annealed again to the oligonucleotide 1 as well as product complementary to oligonucleotide 1 is partially re-anneals to oligonucleotide 1 and partially anneals to oligonucleotide 2 due to an overlapping sequence region between oligonucleotide 1 and oligonucleotide 2. Extension of Pr1 leads to production of additional amount of Pr1 extension product (complementary oligonucleotide 1). The annealing of the complementary  
10 oligonucleotide 1 to oligonucleotide 2 followed by its extension leads to product complementary to oligonucleotides 1+2 (called complimentary oligonucleotides 1+2). Similarly, at step 3 of the process melting, re-annealing and extension lead to the same products as at step 2 plus a product complementary to oligonucleotides 1+2+3. These cycles of melting, annealing and extension are repeated until full-length polynucleotide is formed.  
15 The number of cycles should be equal or more than the number of oligos on microarray device. After formation, the final target polynucleotide molecule is amplified by a PCR process with two primers complementary to the ends of this molecule to the desirable amounts.

In a second embodiment, a plurality of oligonucleotides that together comprise (with overlapping regions) the target polynucleotide sequence are synthesized on a microarray  
20 device (or can be synthesized on beads as a solid substrate), wherein each oligonucleotide sequence further comprises flanking short sequence regions, wherein each flanking sequence region comprises one or a plurality of sequence sites for restriction endonuclease, preferably endonuclease class II (E<sub>RII</sub>) enzymes. Each oligonucleotide is amplified by PCR using appropriate oligonucleotide primers to the flanking sequence regions to form a preparation of a  
25 plurality of oligonucleotides. The preparation of oligonucleotides is treated then with appropriate E<sub>RII</sub> enzyme(s) (specific to the restriction sequences in the flanking sequence regions) to produce flanking fragments and overlapping oligonucleotides that, together comprise the desired polynucleotide sequence. Flanking fragments and PCR primers are removed from the mixture, if desired, by different methods based on size or specific labeling of  
30 the PCR primers. The oligonucleotides resembling the desired target polynucleotide then assembled into the final target polynucleotide molecule using repetition of the primer extension method and PCR amplification of the final molecule.

Specifically, in second embodiment, the assembly process initially uses oligonucleotides immobilized on a microarray device or beads, via immobilization techniques,  
35 such as spotting or ink-jet printing or by direct *in situ* synthesis of the microarray device using various techniques, such as photolithography or electrochemical synthesis. The overlapping oligonucleotide sequences are designed having an overlapping region and one or two flanking sequence regions comprising a restriction class II recognition site (Figure 2A). The assembled oligonucleotides together comprise the target polynucleotide sequence.

The length of flanking sequences is at least the length of REII recognition site. The flanking sequences are designed to have minimal homology to the specific oligonucleotide sequences regions on the microarray device. The flanking sequences can be the same for each oligonucleotide fragment, or be two or more different sequences. For example, a pair of appropriate primers, called Pr1 and Pr2, was designed to amplify each oligonucleotide on a microarray device (Figure 2) by PCR. Each primer may contain a binding moiety, such as biotin, that does not affect their ability to serve as primers. After PCR amplification the amplified ds copy of each oligonucleotide was present in the reaction mixture. This reaction mixture was treated with the appropriate REII enzyme or enzymes specific for the restriction sites in the flanking sequence regions. The digestion sites for REII were designed, after cleavage, to produce the desired specific oligonucleotide sequence fragments that, when assembled will form the target polynucleotide sequence. As a result of digestion a mixture of specific double stranded (ds) overlapping oligonucleotide sequence fragments resembling the structure of desired target polynucleotide, and ds flanking sequences were formed. If desired, these flanking sequences and residual primers are removed from the mixture using specific absorption through specific moieties introduced in the primers (such as, for example, by absorption on avidin beads for biotin-labeled primers), or based on the size difference of the specific oligos and flanking sequences and primers. The mixture of specific oligonucleotide sequences resembling target gene sequence is used to assemble the final target polynucleotide molecule using repeated cycles of melting, self-annealing and polymerase extension followed by PCR amplification of the final target polynucleotide molecule with appropriate PCR primers designed to amplify. This final PCR amplification step is routinely done in the art and described in, for example, Mullis et al., *Cold Spring Harb. Symp. Quant. Biol.* 51 Pt 1:263-73, 1986; and Saiki et al., *Science* 239:487-91, 1988. PCR amplification steps generally follow manufacturers instructions. Briefly, A process for amplifying any target nucleic acid sequence contained in a nucleic acid or mixture thereof comprises treating separate complementary strands of the nucleic acid with a molar excess of two oligonucleotide primers and extending the primers with a thermostable enzyme to form complementary primer extension products which act as templates for synthesizing the desired nucleic acid sequence. The amplified sequence can be readily detected. The steps of the reaction can be repeated as often as desired and involve temperature cycling to effect hybridization, promotion of activity of the enzyme, and denaturation of the hybrids formed.

In another embodiment for the assembly step, oligonucleotide sequences that together comprise the target polynucleotide molecule, are assembled using a ligase chain reaction as described in Au et al., *Biochem. Biophys. Res. Commun.* 248:200-3, 1998. Briefly, short oligonucleotides are joined through ligase chain reaction (LCR) in high stringency conditions to make "unit fragments" (Fifty microliters of reaction mixture contained 2.2 mM of each oligo, 8 units Pfu DNA ligase (Stratagene La Jolla, CA) and reaction buffer provided with the enzyme. LCR was conducted as follows: 95 °C 1 min; 55 °C 1.5 min, 70 °C 1.5 min, 95 °C 30

sec for 15 cycles; 55 °C 2 min; 70 °C 2 min, which are then fused to form a full-length gene sequence by polymerase chain reaction.

In another yet embodiment the ds oligonucleotide sequences are assembled after preparation by chain ligation cloning as described in Pachuk et al., *Gene* 243:19-25, 2000; and  
5 U.S. Patent 6,143,527 (the disclosure of which is incorporated by reference herein). Briefly, chain reaction cloning allows ligation of double-stranded DNA molecules by DNA ligases and bridging oligonucleotides. Double-stranded nucleic acid molecules are denatured into single-stranded molecules. The ends of the molecules are brought together by hybridization to a template. The template ensures that the two single-stranded nucleic acid molecules are aligned  
10 correctly. DNA ligase joins the two nucleic acid molecules into a single, larger, composite nucleic acid molecule. The nucleic acid molecules are subsequently denatured so that the composite molecule formed by the ligated nucleic acid molecules and the template cease to hybridize to each. Each composite molecule then serves as a template for orienting unligated, single-stranded nucleic acid molecules. After several cycles, composite nucleic acid molecules  
15 are generated from smaller nucleic acid molecules. A number of applications are disclosed for chain reaction cloning including site-specific ligation of DNA fragments generated by restriction enzyme digestion, DNase digestion, chemical cleavage, enzymatic or chemical synthesis, and PCR amplification.

With regard to the second embodiment of the inventive process (illustrated in Figure 2),  
20 a target polynucleotide gene sequence (either strand) is divided into number of overlapping oligonucleotide sequences by hand or with a software program, as shown in Figure 1. These oligonucleotide sequences, plus flanking sequences A and B (having one or a plurality of restriction enzyme sites in the flanking region sequence), are synthesized (*in situ*) on microarray device, or on a bead sold surface using standard *in situ* synthesis techniques, or  
25 spotted (pre-synthesized) onto a microarray device using standard oligonucleotide synthesis procedures with standard spotting (e.g., computer-aided or ink jet printing) techniques. The oligonucleotide sequences are amplified, preferably using a PCR process with a pair of primers (Pr1 and Pr2). The primers are optionally labeled with specific binding moieties, such as biotin. The resulting amplified mixture of different amplified oligonucleotide sequences are  
30 double stranded (ds). The mixture of ds oligonucleotide sequences are treated with an appropriate restriction enzyme, such as an REII restriction enzyme (e.g., Mly I enzyme), to produce mixture of different double stranded (ds) overlapping oligonucleotide sequences that can be assembled into the structure of the desired polynucleotide (gene) and ds flanking sequences. Optionally, the flanking sequences and residual primers are removed from the ds  
35 oligonucleotide sequence mixture, preferably by a process of specific absorption using specific binding moieties introduced in the primers (e.g., biotin), or by a process of size fractionation based on the size differences of the specific oligonucleotide sequences and flanking sequences. The mixture of specific oligonucleotide sequences is assembled, for example, by a process of repeated cycles of melting, self-annealing and polymerase extension followed by PCR

amplification of the final molecule with appropriate PCR primers designed to amplify this complete molecule (*e.g.*, as described in Mullis et al., *Cold Spring Harb. Symp. Quant. Biol.* 51 Pt 1:263-73, 1986; and Saiki et al., *Science* 239:487-91, 1988).

In yet another embodiment of the inventive process (illustrated in Figure 3), the oligonucleotide sequences comprising the target polynucleotide sequence are synthesized on a microarray device or bead solid support, each oligonucleotide having a cleavable linker moiety synthesized within the sequence, such that after synthesis, oligonucleotides can be cleaved from the microarray device into a solution. Examples of appropriate cleavable are shown in Figure 3A. In addition to this method of cleavage, a sequence containing RE enzyme site can be synthesized at the ends of oligonucleotides attached to the microarray device. These oligonucleotides on the microarray device then hybridize with an oligonucleotide complementary to this additional flanking sequence and treated with an RE enzyme specific for the RE enzyme site. This process releases oligonucleotide fragments resembling the structure of the target polynucleotide. This set of oligonucleotides then can be assembled into the final polynucleotide molecule using any one of the methods or combination of the methods of ligation, primer extension and PCR.

In a third embodiment of the inventive process, a plurality of oligonucleotides that can be assembled into a full length polynucleotide are synthesized on a microarray device (or beads having a solid surface) having specific cleavable linker moieties (Figure 3A) or capable of being cleaved from the solid support of the microarray device or beads by a chemical treatment. The net effect is to recreate the functional 3' ends and 5' ends of each specific oligonucleotide sequence. After treatment to cleave them, the oligonucleotides (each having overlapping regions) are released into a mixture and used for full-length polynucleotide gene assembly using any of the gene assembly processes described herein.

Specifically, in the third embodiment and as illustrated in Figure 3, a target polynucleotide sequence is dissected into number of overlapping oligonucleotide sequences by a software program or on paper, but not necessarily physically in a laboratory. These oligonucleotide sequences are physically synthesized on a microarray device. In alternative A, the oligonucleotide sequences are connected to the microarray device through cleavable linker moiety. Chemical cleavage under basic conditions (*e.g.*, through addition of ammonia), at cleavable linker CL recreates the usual 3' end of the specific oligonucleotide sequences 1 through N. Oligonucleotide sequences 1 through N are released into a mixture. The mixture of oligonucleotide sequences is used for polynucleotide assembly.

In alternative B, oligonucleotide sequences are connected to a microarray device through additional flanking sequence regions containing a restriction enzyme (RE) sequence site. A second oligonucleotide fragment, complementary to the flanking sequence, is hybridized to the oligonucleotides on the microarray device. This recreates a ds structure at the flanking sequence region, including the RE recognition site. Digestion of this ds DNA structure with RE enzyme specific to the RE recognition site in the flanking sequence region

will release specific oligonucleotides 1 through N into a mixture solution. The oligonucleotides 1 through N are able to assemble into a polynucleotide molecule in solution.

In another example of alternative B, oligonucleotides that together assemble into the polynucleotide, are synthesized on a microarray device, each having a flanking sequence on the microarray side. The flanking sequence further comprises a restriction endonuclease (RE) recognition site (see Figure 3B). Oligonucleotides complementary to the flanking sequence region are added and hybridized to the oligonucleotides on microarray device. After hybridization a RE (restriction enzyme specific to the RE sequence in the flanking region) is added to the microarray device. Specific oligonucleotide sequences are released from the microarray device as a result of RE digestion into a mixture. The mixture of specific oligonucleotide sequences assembled into the full-length polynucleotide sequence.

#### Example 1

This example illustrates assembly of 172-mer polynucleotide sequence from non-cleavable oligonucleotide sequences synthesized on a microarray device according to the first embodiment inventive process (Figures 4 and 5). Three oligonucleotides (sequences shown in Figure 5) were synthesized *in situ* on a microarray device according to an electrochemical process (see U.S. Patent 6,093,302, the disclosure of which is incorporated by reference herein). The oligonucleotide sequences synthesized were amplified by a PCR reaction with primers X (complementary to the strand of oligo#1) and Z (same strand as oligo#3) (Figure 5). After 45 cycles of PCR using a PCR kit with AmplyGold® enzyme (Applied Biosystems) a correct DNA fragment of 172 bp was synthesized (Figure 4). Its subsequent digestion confirmed the specificity of this enzyme with *HpaII* producing two fragments of 106 bp and 68 bp.

#### Example 2

This example illustrates the second embodiment of the inventive process for preparing oligonucleotides for assembly into full-length polynucleotides by PCR and REII (restriction enzyme) digestion. A single oligonucleotide sequence was synthesized on a microarray device according to the procedure in Example 1 (see Figures 2 and 6). The oligonucleotide sequence further comprised 2 flanking sequences, each having a recognition site for a *MlyI* restriction enzyme. This microarray device was subject to a PCR (25 cycles) reaction with two primers (shown in Figure 7) to produce an amplified PCR fragment mixture. The amplified PCR fragment obtained was digested by *MlyI* restriction enzyme and purified by a PCR purification kit (Qiagen) to produce specific oligonucleotides ready for assembly (Figure 7). Similarly, this specific oligonucleotide was purified from the flanking sequences by absorption of the digestion mixture by Streptavidin-agarose (Sigma).

#### Example 3

This example illustrates the assembly of a 290 bp polynucleotide sequence from 9 oligonucleotide sequences, each having flanking sequences containing a *MlyI* restriction site.



Each of the nine different oligonucleotide sequences was synthesized on a microarray device through an *in situ* electrochemistry process as described in example 1 herein.

The microarray device containing the nine specific oligonucleotide sequences (with flanking sequences as shown in Figure 8) was used for PCR amplification of each oligonucleotide sequence using two primers, Primer 1 and 2, described in Figure 6 to form a mixture of ds oligonucleotide sequences. The primers were complementary to the flanking sequences. The mixture of the amplified ds oligonucleotide sequences was digested by *Mly*I enzyme. Specific ds oligonucleotide sequences were purified and then assembled into the final 290 bp polynucleotide sequence in two steps as described in Figure 2 and shown schematically in Figure 9. At the first step of assembly 20 cycles of melting-annealing-extension were used. The final product was amplified using two primers FP1 and FP2 (Figure 9) in 25 cycles of PCR into a 290 bp polynucleotide DNA.

#### Example 4

This example illustrates the creation of a cDNA polynucleotide sequence capable of coding on expression for fusion protein MIP-GFP-FLAG (Macrophage Inflammation Protein – Green Fluorescence Protein – FLAG peptide) using thirty-eight overlapping oligonucleotide sequences (Figure 10). The 38 oligonucleotides were synthesized on a microarray device using an electrochemical *in situ* synthesis approach, as described in example 1. Each oligonucleotide sequence contained a cleavable linker moiety (see Figure 3A) at their 3' end. After simultaneous deprotection and cleavage of these oligonucleotide sequences by concentrated ammonia, the mixture of oligonucleotide sequences was purified by gel-filtration through the spin column. The purified oligonucleotide sequences were assembled into a polynucleotide by a process shown schematically in Figure 3. The resulting DNA polynucleotide was 965 bp and contained both a T7 RNA-polymerase promoter and a coding sequence for MIP-GFP-FLAG fusion protein. The polynucleotide assembled in this example was used in a standard transcription/translation reaction and produced the appropriate MIP-GFP-FLAG fusion protein. The translated protein was purified from the reaction mixture using anti-FLAG resin (Sigma). The functional protein possessed green fluorescence signal in appropriate blue light. Accordingly, this experiment demonstrated that the inventive gene assembly process provided the correct DNA sequence coding for the functional protein.

I claim:

1. A process for assembling a polynucleotide from a plurality of oligonucleotides comprising:
  - (a) synthesizing or spotting a plurality of oligonucleotide sequences on a  
5 microarray device or bead device having a solid or porous surface, wherein a first oligonucleotide is oligo 1 and a second oligonucleotide is oligo 2 and so on, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein the first oligonucleotide sequence has an overlapping sequence region of from about 10 to  
10 about 50 bases that is the same or substantially the same as a region of a second oligonucleotide sequence, and wherein the second oligonucleotide sequence has an overlapping region with a third oligonucleotide sequence and so on;
  - (b) forming complimentary oligo 1 by extending primer 1, wherein primer 1 is complimentary to oligo 1;
  - (c) disassociating complimentary oligo 1 from oligo 1 and annealing  
15 complimentary oligo 1 to both oligo 1 and to the overlapping region of oligo 2, wherein the annealing of complimentary oligo 1 to oligo 2 serves as a primer for extension for forming complimentary oligo 1+2;
  - (d) repeating the primer extension cycles of step (c) until a full-length polynucleotide is produced; and
  - (e) amplifying the assembled complementary full length polynucleotide to produce  
20 a full length polynucleotide in desired quantities.
2. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 1 wherein the solid or porous surface is in the form of a microarray device.
- 25 3. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 2 wherein the microarray device is a semiconductor device having a plurality of electrodes for synthesizing oligonucleotides *in situ* using electrochemical means to couple and decouple nucleotide bases.
4. The process for assembling a polynucleotide from a plurality of  
30 oligonucleotides of claim 1 wherein the primer extension reaction is conducted through a sequential process of melting, annealing and then extension.
5. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 4 wherein the primer extension reaction is conducted in a thermocycler PCR amplification device using the microarray having the plurality of  
35 oligonucleotides bound thereto.
6. A process for assembling a polynucleotide from a plurality of oligonucleotides comprising:
  - (a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a  
microarray device or bead device each having a solid or porous surface, wherein the plurality

of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence has an overlapping region corresponding to a next oligonucleotide sequence within the sequence and further comprises two flanking sequences, one at the 3' end and the other at the 5' end of each oligonucleotide, wherein each flanking sequence is from  
5 about 7 to about 50 bases and comprising a primer region and a sequence segment having a restriction enzyme cleavable site;

(b) amplifying each oligonucleotide using the primer regions of the flanking sequence to form double stranded (ds) oligonucleotides;

(c) cleaving the oligonucleotide sequences at the restriction enzyme cleavable site;

10 and

(d) assembling the cleaved oligonucleotide sequences through the overlapping regions to form a full length polynucleotide.

7. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 6 wherein the flanking sequence is from about 10 to about 20 bases  
15 in length.

8. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 6 wherein the restriction enzyme cleavable site is a class II endonuclease restriction site sequence capable of being cleaved by its corresponding class II restriction endonuclease enzyme.

20 9. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 8 wherein the restriction endonuclease class II site corresponds to restriction sites for a restriction endonuclease class II enzyme selected from the group consisting of Mly I, BspM I, Bae I, BsaX I, Bsr I, Bmr I, Btr I, Bts I, Fok I, and combinations thereof.

25 10. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 6 wherein the flanking sequence further comprises a binding moiety used to purify cleaved oligonucleotides from flanking sequences.

30 11. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 6 wherein the process further comprises the step of labeling the flanking sequence during the amplification step (b) using primer sequences labeled with binding moieties.

12. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 11 wherein a binding moiety is a small molecule able to be captured, such as biotin captured by avidin or streptavidin, or fluorescein able to be captured by an anti-  
35 fluorescein antibody.

13. A process for assembling a polynucleotide from a plurality of oligonucleotides comprising:

(a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality

of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence has an overlapping region corresponding to a next oligonucleotide sequence within the sequence, and further comprises a sequence segment having a cleavable linker moiety;

5 (b) cleaving the oligonucleotide sequences at the cleavable linker site to cleave each oligonucleotide complex from the microarray or bead solid surface to form a soluble mixture of oligonucleotides, each having an overlapping sequence; and

(c) assembling the oligonucleotide sequences through the overlapping regions to form a full length polynucleotide.

10 14. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 13 wherein the cleavable linker is a chemical composition having a succinate moiety bound to a nucleotide moiety such that cleavage produces a 3'hydroxy nucleotide.

15 15. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 14 wherein the cleavable linker is selected from the group consisting of 5'-dimethoxytrityl-thymidine-3'-succinate, 4-N-benzoyl-5'-dimethoxytrityl-deoxycytidine-3'-succinate, 1-N-benzoyl-5'-dimethoxytrityl-deoxyadenosine-3'-succinate, 2-N-isobutyryl-5'-dimethoxytrityl-deoxyguanosone-3'-succinate, and combinations thereof.

20 16. A process for assembling a polynucleotide from a plurality of oligonucleotides comprising:

(a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence has a flanking region at an end attached to the solid or porous surface, and a specific region designed by dissecting the polynucleotide sequence into a plurality of overlapping oligonucleotides, wherein a first overlapping sequence on a first oligonucleotide corresponds to a second overlapping sequence of a second oligonucleotide, and wherein the flanking sequence comprises a sequence segment having a restriction endonuclease (RE) recognition sequence capable of being cleaved by a corresponding RE enzyme;

30 (b) hybridizing an oligonucleotide sequence complementary to the flanking region to form a double stranded sequence capable of interacting with the corresponding RE enzyme;

(c) digesting the plurality of oligonucleotides to cleave them from the microarray device or beads into a solution; and

35 (d) assembling the oligonucleotide mixture through the overlapping regions to form a full length polynucleotide.

17. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 16 wherein the flanking sequence is from about 10 to about 20 bases in length.

18. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 16 wherein the restriction enzyme cleavable site is a class II endonuclease restriction site sequence capable of being cleaved by its corresponding class II restriction endonuclease enzyme.

5 19. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 18 wherein the restriction endonuclease class II site corresponds to restriction sites for a restriction endonuclease class II enzyme selected from the group consisting of Mly I, BspM I, Bae I, BsaX I, Bsr I, Bmr I, Btr I, Bts I, Fok I, and combinations thereof.

10 20. The process for assembling a polynucleotide from a plurality of oligonucleotides of claim 16 wherein the process further comprises a final step of amplifying the polynucleotide sequence using primers located at both ends of the polynucleotide.

21. A process for creating a mixture of oligonucleotide sequences in solution comprising:

15 (a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence further comprises two flanking sequences, one at the 3' end and the other at the 5' end of each oligonucleotide, wherein each flanking sequence is from about 7 to about 50 bases and comprising a primer region and a sequence segment having a restriction enzyme cleavable site;

(b) amplifying each oligonucleotide using the primer regions of the flanking sequence to form a double stranded (ds) oligonucleotides; and

(c) cleaving the double stranded oligonucleotide sequences at the restriction enzyme cleavable site.

25 22. The process for creating a mixture of oligonucleotide sequences in solution of claim 21 wherein the flanking sequence is from about 10 to about 20 bases in length.

30 23. The process for creating a mixture of oligonucleotide sequences in solution of claim 21 wherein the restriction enzyme cleavable site is a class II endonuclease restriction site sequence capable of being cleaved by its corresponding class II restriction endonuclease enzyme.

24. The process for creating a mixture of oligonucleotide sequences in solution of claim 23 wherein the restriction endonuclease class II site corresponds to restriction sites for a restriction endonuclease class II enzyme selected from the group consisting of Mly I, BspM I, Bae I, BsaX I, Bsr I, Bmr I, Btr I, Bts I, Fok I, and combinations thereof.

35 25. The process for creating a mixture of oligonucleotide sequences in solution of claim 21 wherein the flanking sequence further comprises a binding moiety used to purify cleaved oligonucleotides from flanking sequences.

26. The process for creating a mixture of oligonucleotide sequences in solution of claim 21 wherein the process further comprises the step of labeling the flanking sequence during the amplification step (b) using primer sequences labeled with binding moieties.

27. The process for creating a mixture of oligonucleotide sequences in solution of claim 26 wherein a binding moiety is a small molecule able to be captured, such as biotin captured by avidin or streptavidin, or fluorescein able to be captured by an anti-fluorescein antibody.

28. A process for creating a mixture of oligonucleotide sequences in solution comprising:

10 (a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each oligonucleotide sequence has a sequence segment having a cleavable linker moiety;

(b) cleaving the oligonucleotide sequences at the cleavable linker site to cleave  
15 each oligonucleotide sequence from the microarray or bead solid surface to form a soluble mixture of oligonucleotides.

29. The process for creating a mixture of oligonucleotide sequences in solution of claim 28 wherein the cleavable linker is a chemical composition having a succinate moiety bound to a nucleotide moiety such that cleavage produces a 3'-hydroxy nucleotide.

20 30. The process for creating a mixture of oligonucleotide sequences in solution of claim 29 wherein the cleavable linker is selected from the group consisting of 5'-dimethoxytrityl-thymidine-3'-succinate, 4-N-benzoyl-5'-dimethoxytrityl-deoxycytidine-3'-succinate, 1-N-benzoyl-5'-dimethoxytrityl-deoxyadenosine-3'-succinate, 2-N-isobutyryl-5'-dimethoxytrityl-deoxyguanosone-3'-succinate, and combinations thereof.

25 31. A process for creating a mixture of oligonucleotide sequences in solution comprising:

(a) synthesizing *in situ* or spotting a plurality of oligonucleotide sequences on a microarray device or bead device each having a solid or porous surface, wherein the plurality of oligonucleotide sequences are attached to the solid or porous surface, and wherein each  
30 oligonucleotide sequence has a flanking region at an end attached to the solid or porous surface, and a specific region, wherein the flanking sequence comprises a sequence segment having a restriction endonuclease (RE) recognition sequence capable of being cleaved by a corresponding RE enzyme;

(b) hybridizing an oligonucleotide sequence complementary to the flanking region  
35 to form a double stranded sequence capable of interacting with the corresponding RE enzyme;

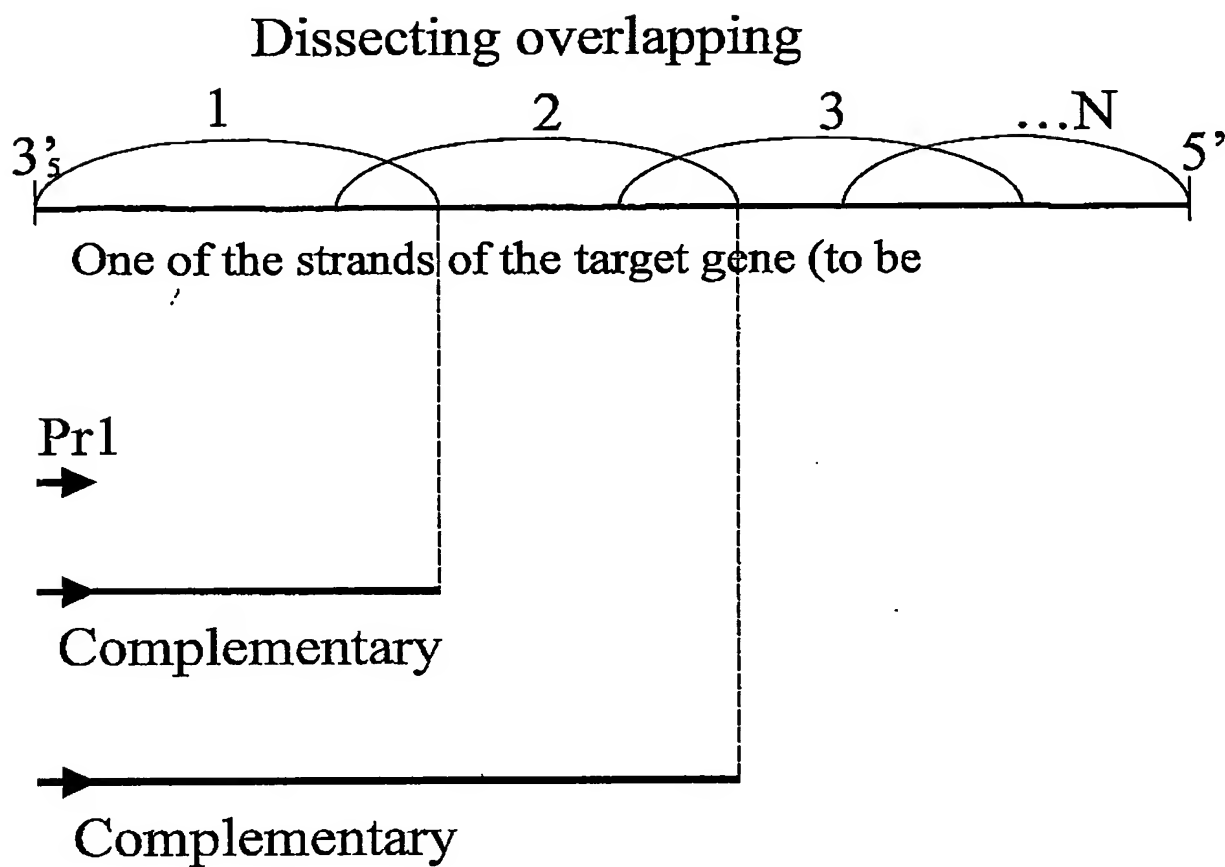
(c) digesting the plurality of oligonucleotides to cleave them from the microarray device or beads into a solution.

32. The process for creating a mixture of oligonucleotide sequences in solution of claim 31 wherein the flanking sequence is from about 10 to about 20 bases in length.

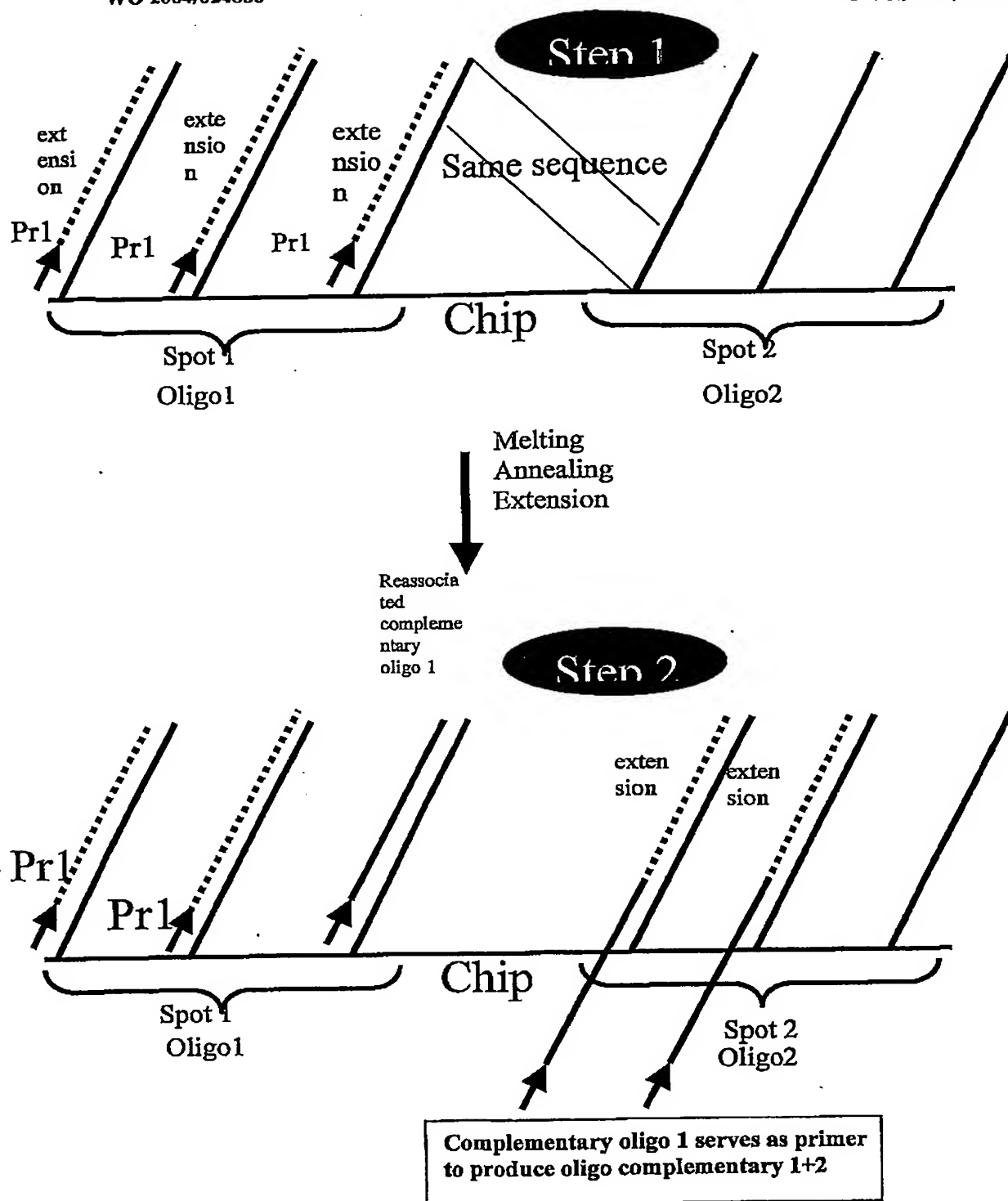
33. The process for creating a mixture of oligonucleotide sequences in solution of claim 31 wherein the restriction enzyme cleavable site is a class II endonuclease restriction site sequence capable of being cleaved by its corresponding class II restriction endonuclease enzyme.

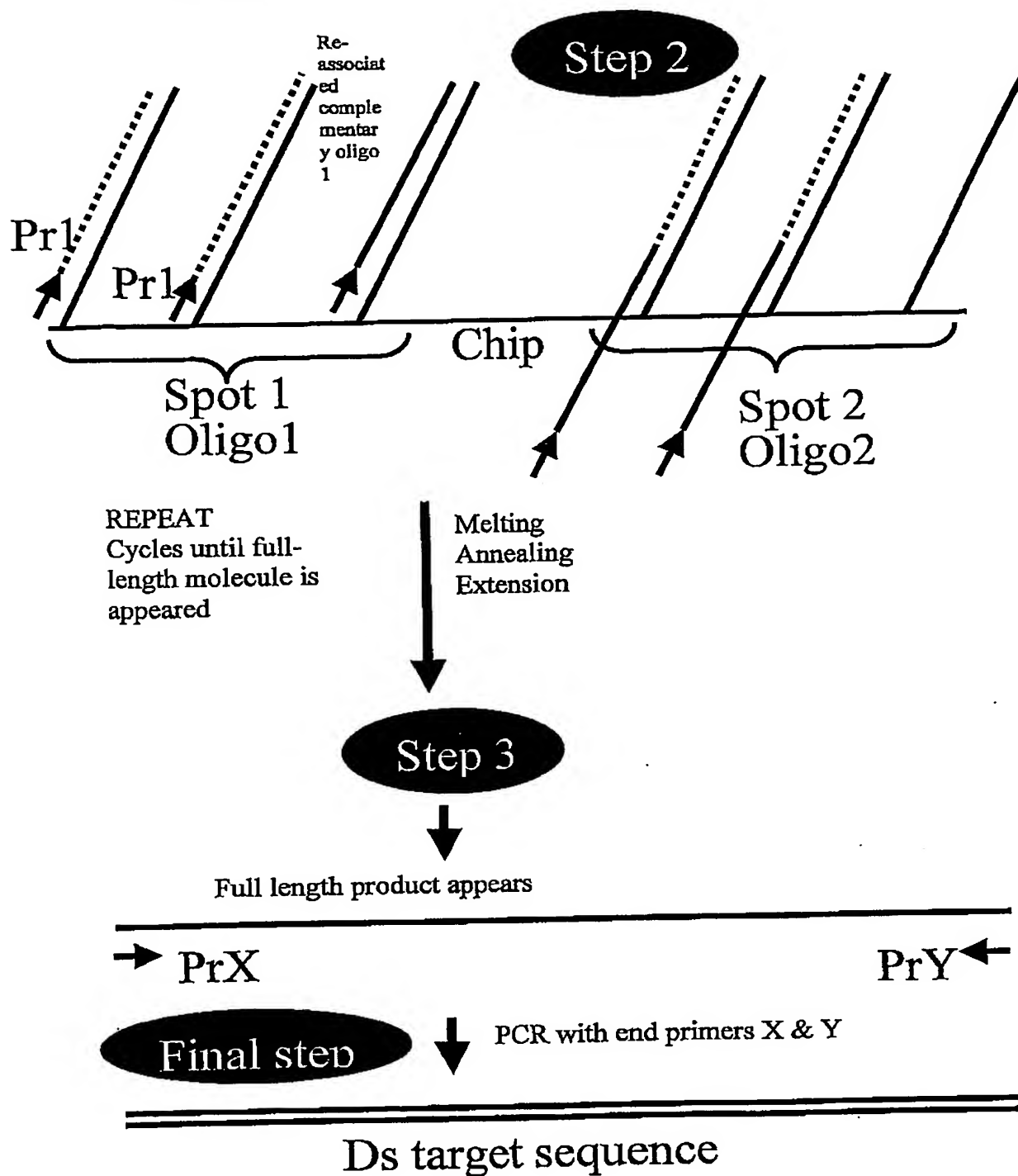
5 34. The process for creating a mixture of oligonucleotide sequences in solution of claim 33 wherein the restriction endonuclease class II site corresponds to restriction sites for a restriction endonuclease class II enzyme selected from the group consisting of Mly I, BspM I, Bae I, BsaX I, Bsr I, Bmr I, Btr I, Bts I, Fok I, and combinations thereof.

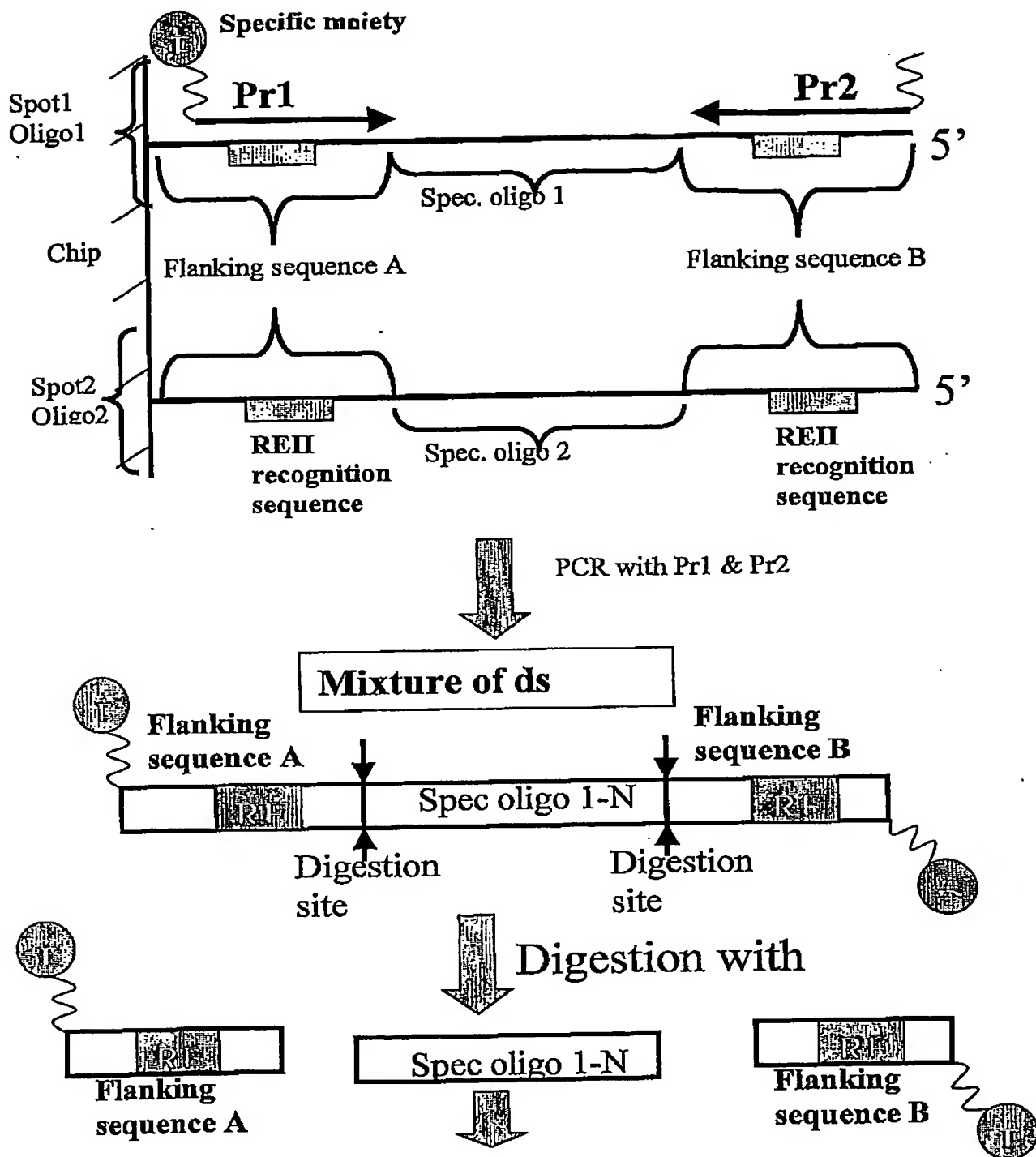
10

**Figure 1A**



**Figure 1B**

**Figure 1C**



Assemble into target sequence by melting, annealing, extension followed by PCR

**Figure 2**

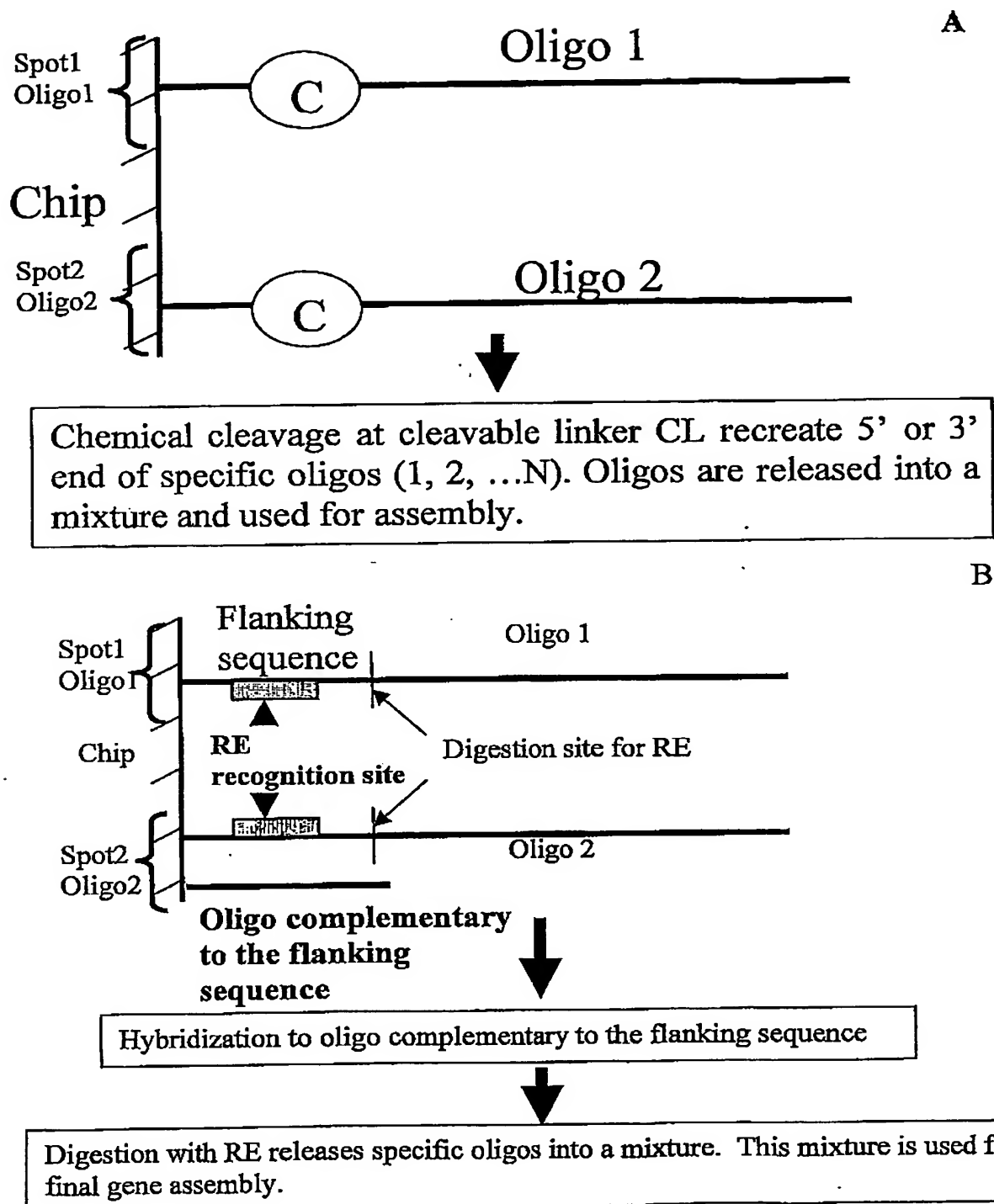
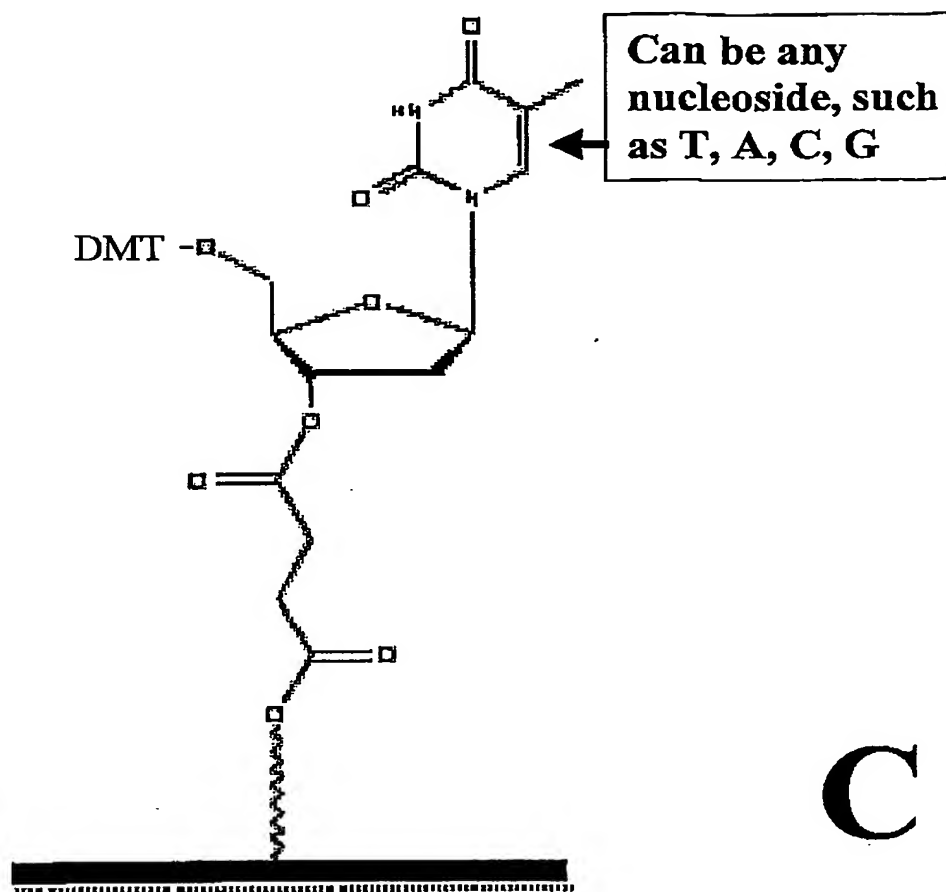
**Figure 3**

Figure 3A

## Cleavable linker for oligonucleotide

5



## On-chip DNA assembly: three non-cleavable oligos

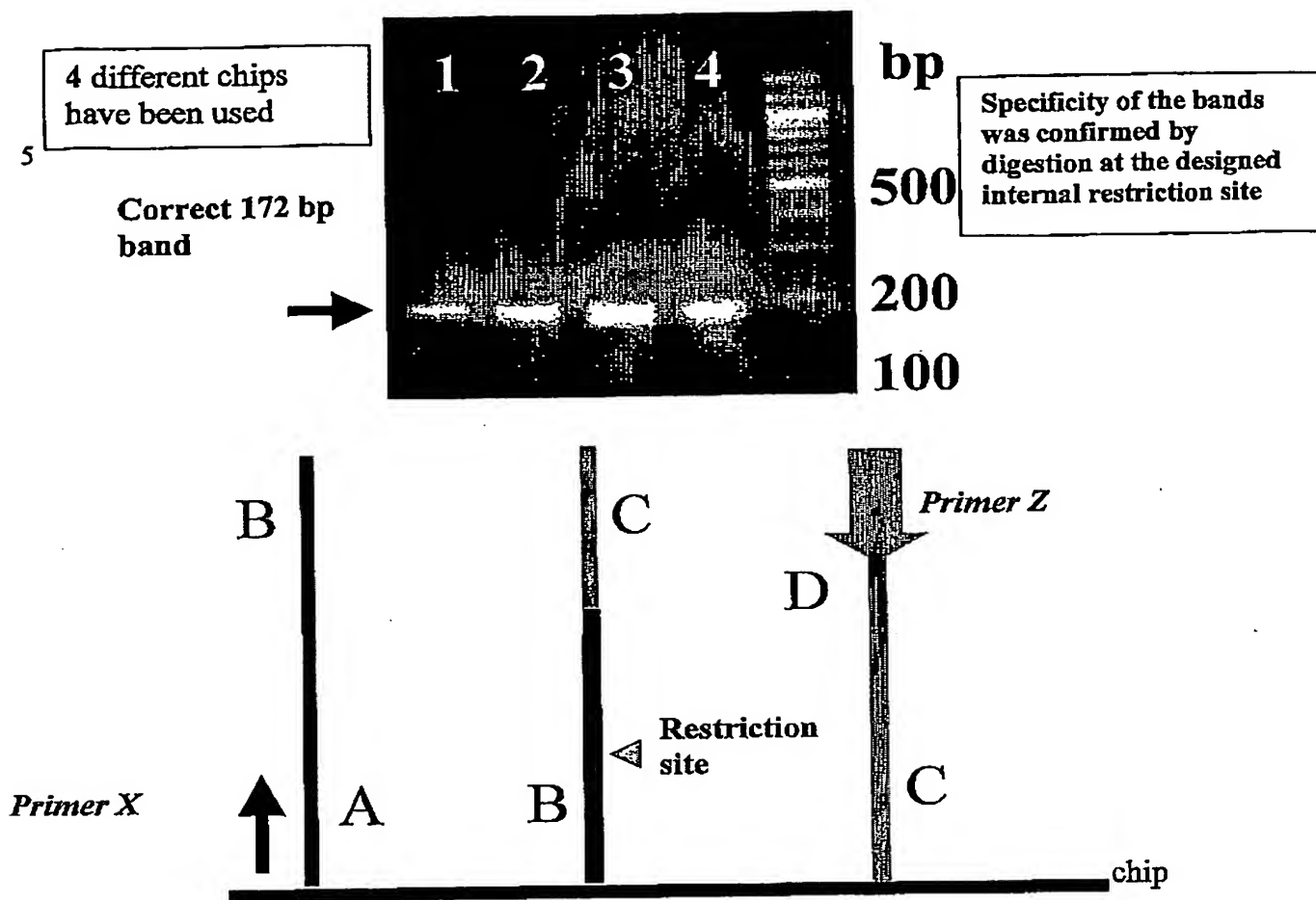


Figure 4

#1  
Chip- Primer X  
3'TAATTATGCTGAgTGATATCCCTTTCTACCTGTGCGGCTGGCGGACGACGAAGT  
5 CGAATGTGGAGGGCCGTCTAAGGTGTCT5' (82mer)

#2  
Chip- HpaII-site  
3'GGACGACGAAGTCGAATGTGGAGGGCCGTCTAAGGTGTCTTAAAGTATCGA  
10 CTGATGAAACTCTGCTCGTCGGTCACGAGGTTC-5' (84mer)

#3  
CHIP-  
15 3'GTATCGACTGATGAAACTCTGCTCGTCGGTCACGAGGTTCCTCGACCACCG  
CATGATGTTTCTGCTACTGCTGTTACGATTATC-5' (86mer)  
Primer Z

Final product: 172mer (one of two strands is shown, direction is 3' to 5' for convenience)  
20 3'TAATTATGCTGAgTGATATCCCTTTCTACCTGTGCGGCTGGCGGACGACGAAGT  
CGAATGTGGAGGGCCGTCTAAGGTGTCTTAAAGTATCGACTGATGAAACTCTG  
CTCGTCGGTCACGAGGTTCCTCGACCACCGCATGATGTTTCTGCTACTGCTGTT  
CACGATTATC5'  
25

Figure 5

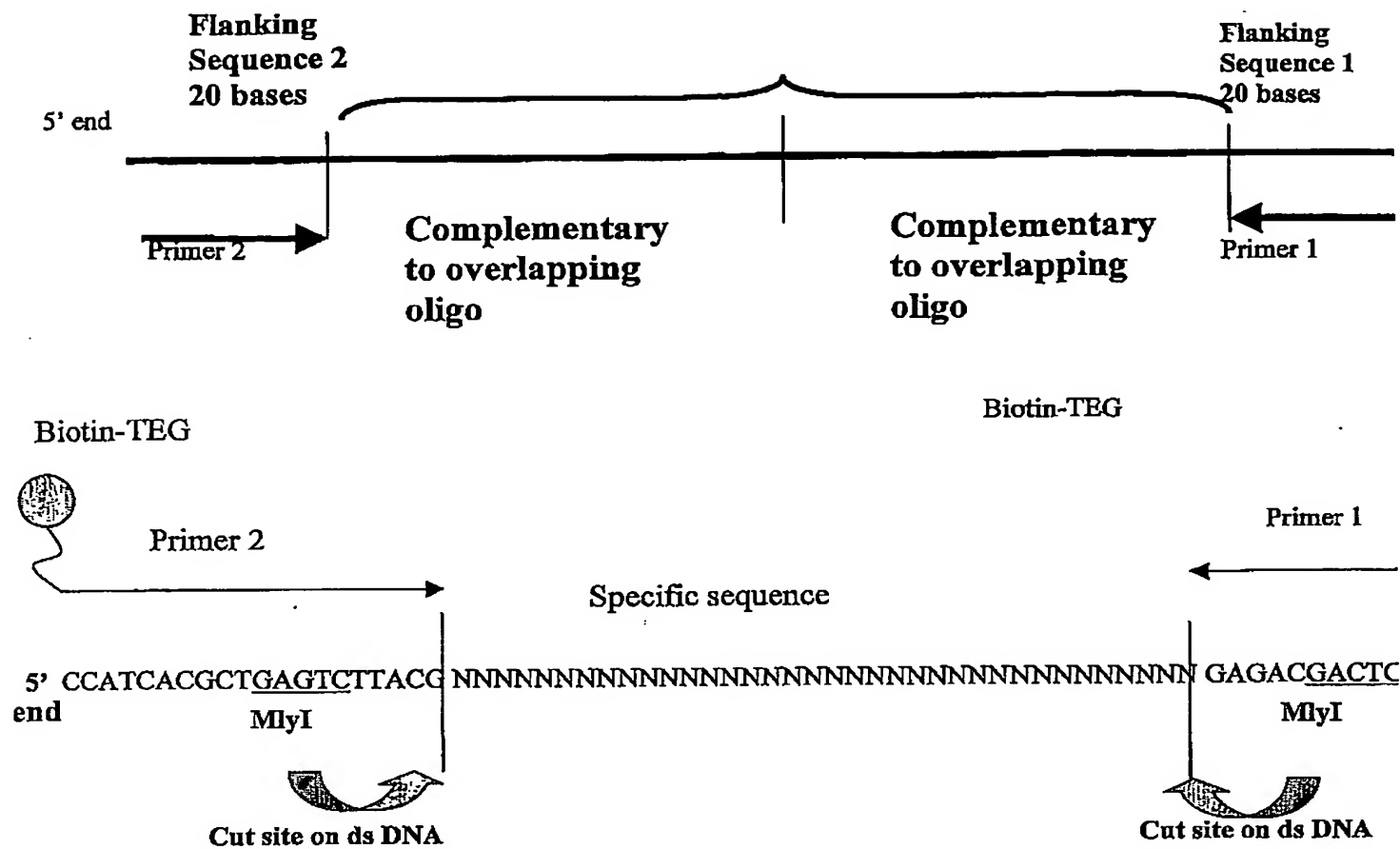


Figure 6



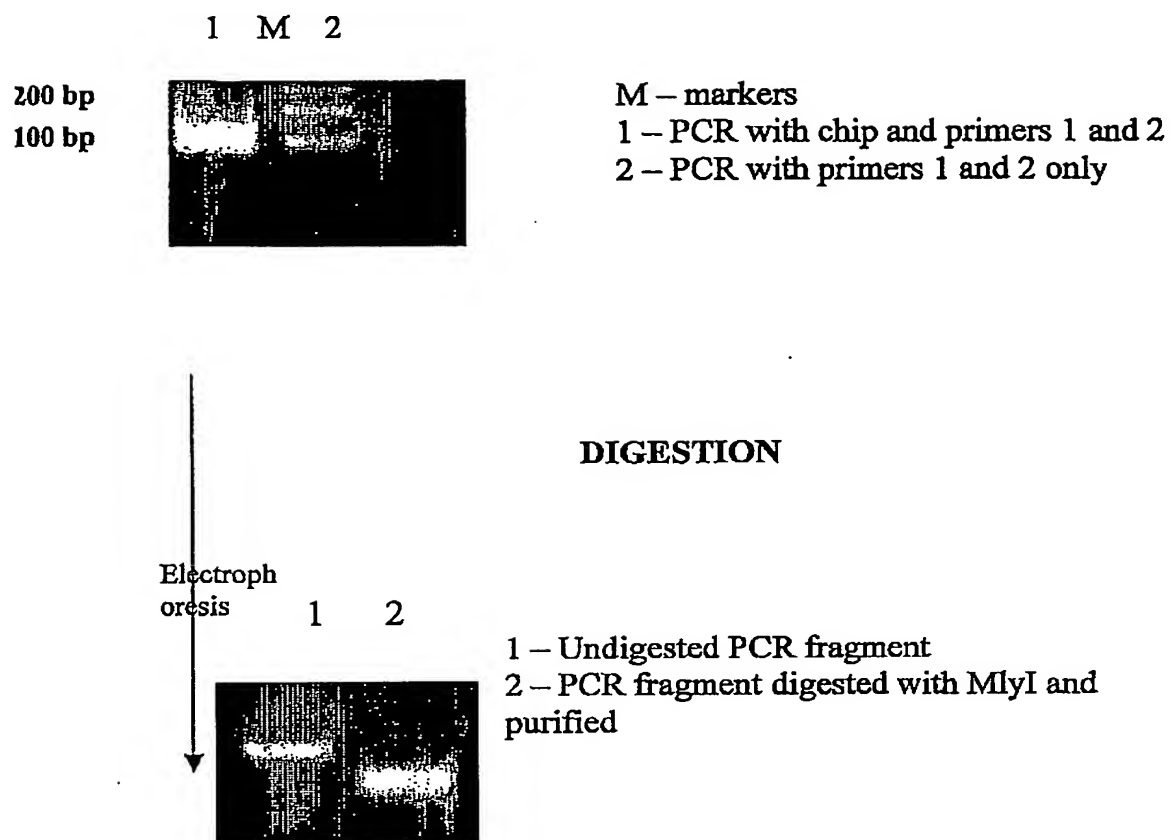


Figure 7

1:  
CCATCACGCTGAGTCTTACGTACGTAATACGACTCACTATAGGGAAAGTCGCCACCATG  
GACACGCCGACGAGACGACTCCTAATCGAA

5 2:  
CCATCACGCTGAGTCTTACGCGCCTGCTGCTCAGCTACACCTCCCGGCAGATTCCACAG  
AATTTGAGACGACTCCTAATCGAA

3:  
CCATCACGCTGAGTCTTACGATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAG  
10 CCCGGTGTCGAGACGACTCCTAATCGAA

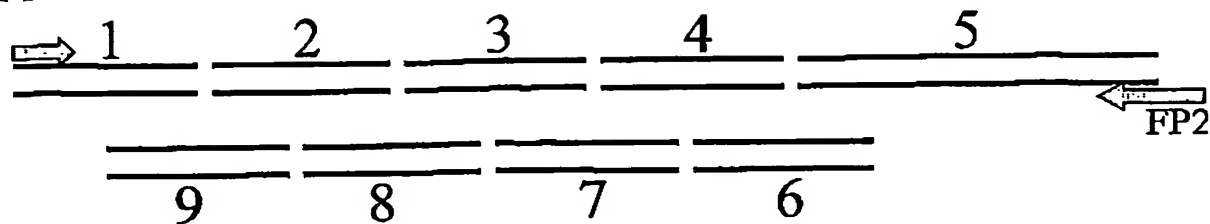
4:  
CCATCACGCTGAGTCTTACGATCTTCCTAACCAAGCGAAGCCGGCAGGTCTGTGCTGACC  
CCGAGACGACTCCTAATCGAA

5:  
15 CCATCACGCTGAGTCTTACGCAGGCACTCAGCTCTACGGGGCCGTCGCCGATGGGGGTG  
TTCTGCTGGTAGTGGTCGGCGAGCTGCATATTTCTGGACCCACTCCTCACTGAGACGACTC  
CTAATCGAAC

6:  
20 CCATCACGCTGAGTCTTACGATATTTCTGGACCCACTCCTCACTGGGGTCAGCACAGACC  
TGCCGAGACGACTCCTAATCGAA

7:CCATCACGCTGAGTCTTACGGGCTTCGCTTGGTTAGGAAGATGACACCGGGCTTGGAG  
CACTGGCGAGACGACTCCTAATCGAA

8:CCATCACGCTGAGTCTTACGTGCTCGTCTCAAAGTAGTCAGCTATGAAATTCTG  
25 TGGAATCTGCCGAGACGACTCCTAATCGAA 9:  
CCATCACGCTGAGTCTTACGGGGAGGTGTAGCTGAAGCAGCAGGCGGTCGGCGTGTCCA  
TGGTGGCGACGAGACGACTCCTAATCGAA



FP1 and FP2 are final primers 1 and 2 used to PCR the target sequence

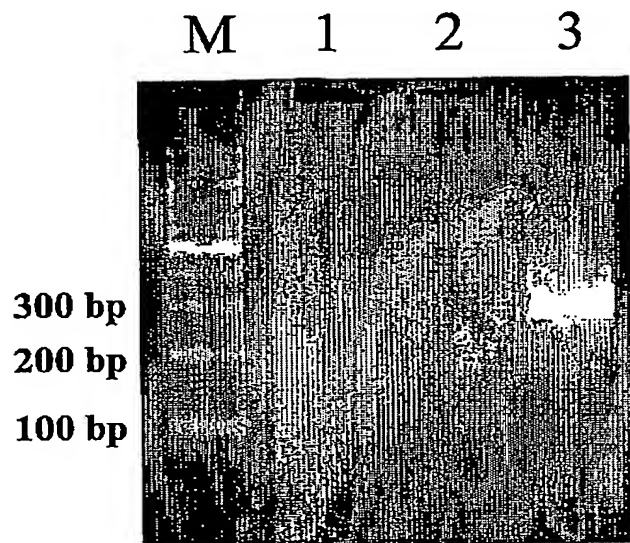


Figure 9

BEST AVAILABLE COPY

1: TACGTAATACGACTCACTATAGGGAAAGTCGCCACCATGGACACGCCGAC  
2: CGCCTGCTGCTTCAGCTACACCTCCCGGCAGATTCCACAGAATTTC  
5 3: ATAGCTGACTACTTTGAGACGAGCAGCCAGTGCTCCAAGCCCGGTGTC  
4: ATCTTCCTAACCAAGCGAAGCCGGCAGGTCTGTGCTGACCCC  
10 5: AGTGAGGAGTGGGTCCAGAAATATGTCAGCGACCTAGAGCTGAGTGC  
6: ATATTTCTGGACCCACTCCTCACTGGGGTCAGCACAGACCTGCC  
7: GGCTTCGCTTGGTAGGAAGATGACACCGGGCTTGAGCACTGGC  
15 8: TGCTCGTCTCAAAGTAGTCAGCTATGAAATTCTGTGGAATCTGCC  
9: GGGAGGTGTAGCTGAAGCAGCAGGCGGTGCGCGTGTCCATGGTGGCGAC  
20 1F: GGTGAACAGCTCCTCGCCCTTGCTCACCATGGCACTCAGCTCTAGGTCGCTGAC  
2F: CATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTC  
3F: TTGTGGCCGTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCC  
25 4F: GAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGC  
5F: TTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAAC  
30 6F: GAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACC  
7F: CAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGC  
8F: ACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGC  
35 9F: GGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTGAGGGTGGTCACGAGGGTGGGC  
10F: GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTC  
40 11F: GTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTGGTC  
12F: TTCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTC  
13F: GGGTCTTGTAGTTGCCGTGCTCCTTGAAGAAGATGGTGCCTCCTGGAC  
45 14F: AAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGC  
15F: AGCTCGATGCGGTTCAGGAGGTTCGCCCTCGAACTTCACCTCGGCGC  
50 16F: GACACCCTGGTGAACTGAGGAGGATGAAGGGCATCGACTTCAAGGAGGAC

Figure 10

(continued)

17F: TCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTC  
18F: GGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTC  
5 19F: GTTCTTCTGCTTGTGCGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTAC  
20F: TATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATC  
10 21F: ACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCC  
22F: CGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGC  
23F: ACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTGC  
15 24F: AGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACC  
25F: TTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGTTGTGCGGCAGCAGC  
20 26F: TGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCAC  
27F: GGCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTC  
28F: ATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGAC  
25 29F: GGCGGCCGCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGC

---

Figure 10  
(continuation)

## Sequence Listing

<110> Oleinikov Andrew V.  
<120> MICROARRAY SYNTHESIS AND ASSEMBLY OF GENE-  
LENGTH POLYNUCLEOTIDES  
<130> 0601WO  
<140> 10/243,367  
<141> 2003-09-12  
<160> 51

<210> 1  
<211> 82  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Primer X #1 Figure 5  
<400> 1

taattatgct gagtgatatc cctttctacc tgtgcggctg gcggacgacg  
aagtcgaatg tggagggccg tctaaggtgt ct 82

<210> 2  
<211> 84  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Primer X #2 Figure 5  
<400> 2

ggacgacgaa gtcgaaatgtg gagggccgctc taaggtgtct taaagtatcg  
actgatgaaa ctctgctcgt cggtcacgag gttc 84

<210> 3  
<211> 86  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Primer Z #3 Figure 5  
<400> 3

gtatcgactg atgaaactct gctcgtcggt cacgaggttc cctcgaccac  
cgcgatgatgt ttctgctact gctgttcacg attatc 86

<210> 4  
<211> 172  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Final assembled product Figure 5  
<400> 4

taattatgct gagtgatatc cttttctacc tgtgcggctg gcggacgacg  
aagtcgaatg tggagggccg tctaagggtg cttaaagtat cgactgatga  
aactctgctc gtcggtcacg aggttccttc gaccaccgca tgatgtttct  
gctactgctg ttcaogatta tc 172

<210> 5  
<211> 89  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #1 Figure 8  
<400> 5

ccatcacgct gagtcttacg tacgtaatac gactcactat agggaaagtc  
gccaccatgg acacgccgac gagacgactc ctaatcgaa 89

<210> 6  
<211> 85  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #2 Figure 8  
<400> 6

ccatcacgct gagtcttacg cgcttctgctg ttcagctaca cctccccggca  
gattccacag aatttcgaga cgactcctaa tcgaa 85

<210> 7  
<211> 87  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #3 Figure 8  
<400> 7

ccatcacgct gagtcttacg atagctgact actttgagac gagcagccag  
tgctccaagc ccggtgtcga gacgactcct aatcgaa 87

<210> 8  
<211> 81  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #4 Figure 8  
<400> 8

ccatcacgct gagtcttacg atcttcctaa ccaagcgaag ccggcaggtc  
tgtgtgacc ccgagacgac tcctaatacga a 81

<210> 9  
<211> 130  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #5 Figure 8  
<400> 9

ccatcacgct gagtcttacg caggcactca gctctacggg gccgtcgccg  
atgggggtgt tctgctggta gtggcggcg agctgcataat ttctggaccc  
actcctcact gagaagactc ctaatcgaac 130

<210> 10  
<211> 83  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #6 Figure 8  
<400> 10

ccatcacgct gagtcttacg atatttctgg acccactcct cactggggtc  
agcacagacc tgccgagacg actcctaatac gaa 83

<210> 11  
<211> 84  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #7 Figure 8  
<400> 11

ccatcacgct gagtcttacg ggcttcgctt ggtaggaag atgacaccgg  
gcttgagaca ctggcgagac gactcctaata cgaa 84

<210> 12  
<211> 84  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #8 Figure 8  
<400> 12

ccatcacgct gagtcttacg tgctcgtctc aaagtagtca gctatgaaat  
tctgtggaat ctggcgagac gactcctaata cgaa 84

<210> 13  
<211> 88  
<212> DNA



<213> Artificial Sequence  
<220>  
<223> Fragment #9 Figure 8  
<400> 13

ccatcacgct gagtcttacg gggagggtgta gctgaagcag caggcgggtcg  
gcgtgtccat ggtggcgacg agacgactcc taatcgaa 88

<210> 14  
<211> 50  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #1 Figure 10  
<400> 14

tacgtaatac gactcactat agggaaagtc gccaccatgg acacgccgac 50

<210> 15  
<211> 46  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #2 Figure 10  
<400> 15

cgctgtgtgc ttcagctaca cctcccggca gattccacag aatttc 46

<210> 16  
<211> 48  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #3 Figure 10  
<400> 16

atagctgact actttgagac gaggagccag tgctccaagc ccggtgtc 48

<210> 17  
<211> 42  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #4 Figure 10  
<400> 17

atcttcctaa ccaagcgaag ccggcaggtc tgtgtgacc cc 42

<210> 18  
<211> 47  
<212> DNA

<213> Artificial Sequence  
<220>  
<223> Fragment #5 Figure 10  
<400> 18

agtgaggagt gggccagaa atatgtcagc gacctagagc tgagtgc 47

<210> 19  
<211> 44  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #6 Figure 10  
<400> 19

atatttctgg acccactcct cactggggtc agcacagacc tgcc 44

<210> 20  
<211> 45  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #7 Figure 10  
<400> 20

ggcttcgctt ggtaggaag atgacaccgg gcttgagca ctggc 45

<210> 21  
<211> 45  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #8 Figure 10  
<400> 21

tgctcgtctc aaagtagtca gctatgaaat tctgtggaat ctgcc 45

<210> 22  
<211> 49  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #9 Figure 10  
<400> 22

gggaggtgta gctgaagcag caggcggtcg gcgtgtccat ggtggcgac 49

<210> 23  
<211> 54  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Fragment #1F Figure 10  
<400> 23

ggtgaacagc tcctcgccct tgctcaccat ggcactcagc tctaggtegc  
tgac 54

<210> 24  
<211> 52  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #2F Figure 10  
<400> 24

catgggtgagc aagggcgagg agctgttcac cgggggtggtg cccatcctgg tc 52

<210> 25  
<211> 50  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #3F Figure 10  
<400> 25

ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg gcaccacccc 50

<210> 26  
<211> 51  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #4F Figure 10  
<400> 26

gagctggacg gcgacgtaaa cggccacaag ttcagcgtgt ccggcgaggg c 51

<210> 27  
<211> 45  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #5F Figure 10  
<400> 27

ttgcggtagg tggcatcgcc ctcgccctcg ccggacacgc tgaac 45

<210> 28  
<211> 48  
<212> DNA  
<213> Artificial Sequence

<220>  
 <223> Fragment #6F Figure 10  
 <400> 28

gagggcgatg ccacctacgg caagctgacc ctgaagttca tctgcacc 48

<210> 29  
 <211> 48  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Fragment #7F Figure 10  
 <400> 29

cagggcacgg gcagcttgcc ggtggtgcag atgaacttca gggtcagg 48

<210> 30  
 <211> 54  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Fragment #8F Figure 10  
 <400> 30

accggcaagc tgcccgtgcc ctggcccacc ctcgtgacca ccctgaccta  
 cggc 54

<210> 31  
 <211> 55  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Fragment #9F Figure 10  
 <400> 31

ggggtagcgg ctgaagcact gcacgccgta ggtcaggggtg gtcacgaggg  
 tgggc 55

<210> 32  
 <211> 48  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Fragment #10F Figure 10  
 <400> 32

gtgcagtgct tcagccgcta ccccgaccac atgaagcagc acgacttc 48

<210> 33  
 <211> 51  
 <212> DNA

<213> Artificial Sequence  
<220>  
<223> Fragment #11F Figure 10  
<400> 33  
  
gtagccttcg ggcattggcgg acttgaagaa gtcgtgctgc ttcattgtggt c 51  
  
<210> 34  
<211> 51  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #12F Figure 10  
<400> 34  
  
ttcaagtccg ccatgcccga aggctacgtc caggagcgca ccattcttct c 51  
  
<210> 35  
<211> 49  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #13F Figure 10  
<400> 35  
  
gggtcttgta gttgccgtcg tccttgaaga agatgggtgcg ctctctggac 49  
  
<210> 36  
<211> 48  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #14F Figure 10  
<400> 36  
  
aaggacgacg gcaactacaa gacccgcgoc gaggtgaagt tcgagggc 48  
  
<210> 37  
<211> 49  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #15F Figure 10  
<400> 37  
  
agctcgatgc ggttcaccag ggtgtcgccc tcgaacttca cctcggcgc 49  
  
<210> 38  
<211> 51  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Fragment #16F Figure 10  
<400> 38  
gacaccctgg tgaaccgcat cgagctgaag ggcacgcact tcaaggagga c 51  
<210> 39  
<211> 51  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #17F Figure 10  
<400> 39  
tccagcttgt gccccaggat gttgccgtcc tccttgaagt cgatgccctt c 51  
<210> 40  
<211> 51  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #18F Figure 10  
<400> 40  
ggcaacatcc tggggcacaa gctggagtac aactacaaca gccacaacgt c 51  
<210> 41  
<211> 52  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #19F Figure 10  
<400> 41  
gttcttctgc ttgtcggcca tgatatagac gttgtggctg ttgtagttgt ac 52  
<210> 42  
<211> 51  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #20F Figure 10  
<400> 42  
tatatcatgg ccgacaagca gaagaacggc atcaagggtga acttcaagat c 51  
<210> 43  
<211> 50  
<212> DNA  
<213> Artificial Sequence  
<220>

<223> Fragment #21F Figure 10  
<400> 43

acgctgccgt cctcgatggt gtggcggatc ttgaagttca ccttgatgcc 50

<210> 44  
<211> 49  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #22F Figure 10  
<400> 44

cgccacaaca tcgaggacgg cagcgtgcag ctgcccgacc actaccagc 49

<210> 45  
<211> 51  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #23F Figure 10  
<400> 45

acggggccgt cgccgatggg ggtgttctgc tggtagtggc cggcgagctg c 51

<210> 46  
<211> 51  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #24F Figure 10  
<400> 46

agaacacccc catcggcgac ggccccgtgc tgctgcccga caaccactac c 51

<210> 47  
<211> 49  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #25F Figure 10  
<400> 47

tttgctcagg gcggaactggg tgctcaggta gtggttgctg ggcagcagc 49

<210> 48  
<211> 50  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Fragment #26F Figure 10

<400> 48

tgagcaccca gtccgccctg agcaaagacc ccaacgagaa gcgcgatcac 50

<210> 49

<211> 51

<212> DNA

<213> Artificial Sequence

<220>

<223> Fragment #27F Figure 10

<400> 49

ggcggtcacg aactccagca ggaccatgtg atcgcgcttc tcgttggggc c 51

<210> 50

<211> 51

<212> DNA

<213> Artificial Sequence

<220>

<223> Fragment #28F Figure 10

<400> 50

atgggtcctgc tggagttcgt gaccgcccgc gggatcactc tcggcatgga c 51

<210> 51

<211> 49

<212> DNA

<213> Artificial Sequence

<220>

<223> Fragment #29F Figure 10

<400> 51

ggcggccgct ttacttgtag agctcgtcca tgccgagagt gatccccgc 49